

65 35000 (5)

BEST AVAILABLE COPY

(12) NACH DEM VERTRÄG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro(43) Internationales Veröffentlichungsdatum
29. November 2001 (29.11.2001)

PCT

(10) Internationale Veröffentlichungsnummer
WO 01/90932 A2

(51) Internationale Patentklassifikation: **G06F 17/30** (74) Anwalt: **BARTH, Stephan; Reinhard, Skuhra, Weise & Partner GbR, Postfach 44 01 51, 80750 München (DE).**

(21) Internationales Aktenzeichen: **PC1/EP01/05746**

(22) Internationales Anmeldedatum: **18. Mai 2001 (18.05.2001)**

(25) Einreichungssprache: **Deutsch**

(26) Veröffentlichungssprache: **Deutsch**

(30) Angaben zur Priorität: **100 24 733.4 19. Mai 2000 (19.05.2000) DE**

(71) Anmelder und
(72) Erfinder: **SPEHR, Clemente [DE/DE]; Holzstr. 33, 80469 München (DE).**

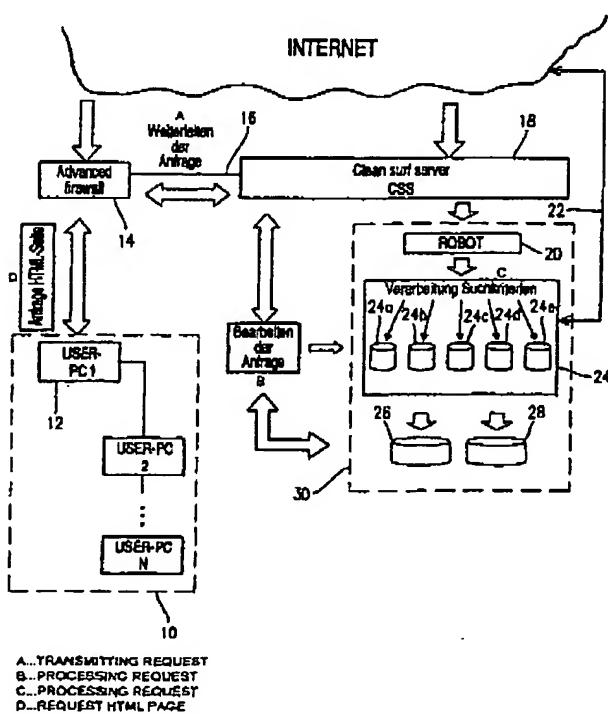
(81) Bestimmungsstaaten (national): **AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SH, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, YA, ZW.**

(84) Bestimmungsstaaten (regional): **ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).**

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD AND DEVICES FOR BLOCKING DATA WHICH CAN BE REQUESTED FROM A NETWORK

(54) Bezeichnung: VERFAHREN UND VORRICHTUNG ZUM ABBLOCKEN VON AUS EINEM NETZWERK ANFORDERBAREN DATEN



(57) Abstract: The invention relates to a method and devices for blocking resources which can be requested from a network and which have an undesirable content, and to a method for filtering data of this type out of any quantity of data. Requested web pages are e.g. checked for integrity before being used by the end user. The information resources that can be requested are essentially fully automatically checked and qualified by an inventive filtering program based on the detection and weighted evaluation of individual suspicious information elements.

(57) Zusammenfassung: Ein Verfahren und Vorrichtungen zum Abblocken von aus einem Netzwerk anforderbaren Ressourcen mit unerwünschtem Inhalt, sowie ein Verfahren zum Ausfiltern solcher Daten aus einer beliebig grossen Datenmenge wird vorgeschlagen. Angeforderte Webseiten werden beispielsweise vor dem Verbrauch beim Endbenutzer auf ihre Integrität hin kontrolliert. Die anforderbaren Informationsressourcen werden im wesentlichen vollautomatisch durch ein erfindungsmässiges Filterprogramm beruhend auf Erkennung und gewichteter Bewertung einzelner "verdächtiger" Informationselementen untersucht und qualifiziert.

WO 01/90932 A2

WO 01/90932 A2



OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML,
MR, NE, SN, TD, TG).

Zur Erklärung der Zweibuchstaben-Codes und der anderen
Abkürzungen wird auf die Erklärungen ("Guidance Notes on
Codes and Abbreviations") am Anfang jeder regulären Ausgabe
der PCT-Gazette verwiesen.

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu
veröffentlichen nach Erhalt des Berichts

WO 01/90932

PCT/EP01/05746

Verfahren und Vorrichtung zum Abblocken von aus einem
Netzwerk anforderbaren Daten

5

STAND DER TECHNIK

Die vorliegende Erfindung betrifft ein Verfahren und eine Vorrichtung zum Abblocken von aus einem Netzwerk anforderbaren Daten mit unerwünschtem Inhalt sowie ein Verfahren zum Ausfiltern solcher Daten aus einer beliebig großen Datenmenge.

Obwohl auf beliebige Daten und Netzwerke anwendbar, werden die vorliegende Erfindung sowie die ihr zugrundeliegende Problematik in Bezug auf eine Situation erläutert, in der mehrere, miteinander vernetzte Endbenutzer-PCs einen Firewall Zugang zum Internet besitzen. Das Internet wird heutzutage von vielen Menschen genutzt. Es ist eine hervorragende, komplexe, unter Umständen einfach zu rech�chierende Informationsressource für eine Vielfalt von Ressourcen, wie beispielsweise Texte, Bilder, Daten, Tonsequenzen oder Bildsequenzen beziehungsweise Kombinationen daraus, wie es beispielsweise in Videofilmen der Fall ist. Die Informationen werden dabei grundsätzlich auf Anforderung eines Endbenutzers von einer meist für die angeforderte Information spezifischen Datenquelle, einem Netzserver, zum Endbenutzer transportiert, wo sie entwe-

WO 01/90932

PCT/EP01/05746

- 2 -

der nur dargestellt oder auch gespeichert und weiterverarbeitet werden können.

5 Mit zunehmender Akzeptanz des Internets in weiten Kreisen von Wirtschaft und Bevölkerung haben sich die Inhalte der im Internet dargebotenen Webseiten stark diversifiziert: einerseits gibt es informative, sehr nützliche Informationen die für den Verbrauch beim Endbenutzer unproblematisch sind. Andererseits gibt es jedoch eine Vielzahl 10 nicht gewünschter Informationen, beispielsweise Informationen, die thematisch im Bereich Sex, Hardcore, Kinderporno, Gewalt, oder auch reiner Werbung liegen.

15 Je nach Alter, Interessen und Hintergrund eines oder mehrerer Endbenutzer dieser Informationen ist es wünschenswert, bestimmte oder alle im Netz verfügbaren Informationen, die zu einem oder mehreren der vorgenannten Themen eindeutig zuzuordnen sind, auszufiltern, um den Endbenutzer nicht damit zu belasten.

20 Solche Informationen werden im folgenden als nicht gewünschte Informationen bezeichnet.

25 Es gibt freilich gewisse objektive Maßstäbe nach denen solche Informationen gefiltert werden könnten. Darunter fallen beispielsweise Informationen, in denen die Gewalt verherrlicht wird, die Kinderpornos zeigen, oder gewisse, moralisch stark bedenkliche Sex and Crime-Inhalte, um nur die wichtigsten zu nennen. Für solche Inhalte erscheint

WO 01/90932

PCT/EP01/05746

- 3 -

eine globale, klassische Indizierung angebracht. Das Problem dabei ist jedoch, wie solche Informationen, beispielsweise in Form von Webseiten wirksam vor dem Endbenutzer abgeblockt werden können. Ein weiterer Aspekt, der 5 bei jenem Abblocken zu berücksichtigen ist, ist die Tatsache, daß ein und der selbe Inhalt nicht für alle Menschen gleich schädlich beziehungsweise unerwünscht ist. So gibt es beispielsweise Menschen, die sich durch Werbeblöcke oder Werbebanner kaum stören lassen oder aber Menschen, die auf die Einblendung solcher Werbemittel sehr 10 sensibel reagieren. Auch kann ein erwachsener Mensch ein größeres Maß an den oben genannten Sex and Crime-Informationen verarbeiten, ohne daran Schaden zu nehmen, im Vergleich zu einem Kind. Da Kinder jedoch in zunehmendem Maße auch als Endbenutzer in Frage kommen, müssen 15 Kinder beispielsweise selektiv vor unerwünschten oder verbotenen Inhalten geschützt werden.

20 Im Stand der Technik befindliche Möglichkeiten, selektiv bestimmte Inhalte dem Endbenutzer vorzuenthalten, sind nur über Texterkennung gegeben. Insbesondere kann dem Endbenutzer der Zugang nur zum PC als Maschine über ein Passwort möglich sein, wobei das Passwort auch beim Aufruf eines Browser-Programms verlangt werden kann.

25 Dies hat jedoch den Nachteil, daß der Endbenutzer auch keine für ihn nützlichen oder wertvollen Daten aus dem Netz ziehen kann.

WO 01/90932

PCT/EP01/05746

- 4 -

VORTEILE DER ERFINDUNG

Das erfindungsgemäße Verfahren mit den Merkmalen des Anspruchs 1, das Verfahren mit den Merkmalen des Anspruchs 5 sowie die entsprechenden Vorrichtungen gemäß Ansprüchen 16 bis 19 weisen gegenüber den bekannten Lösungsansätzen den Vorteil auf, daß die Zugangskontrolle flexibler als bisher gestaltbar ist. Fordert der Endbenutzer beispielsweise aus dem Internet bestimmte Daten an, so werden diese 10 Daten vor einer Darstellung beim Endbenutzer daraufhin untersucht, ob sie bestimmten, flexibel bestimmmbaren Filterkriterien genügen oder nicht. Diese Filterkriterien werden dann als Basis dafür herangezogen, zu entscheiden, ob die Daten als 'ungewünscht' vom Endbenutzer abzublocken 15 sind, oder ob sie als tolerabel dem Endbenutzer zu-führbar sind. Es wird also ein sogenannter Filterserver zwischen den Endbenutzer-PC und dem Informationsnetzwerk geschaltet, der vorzugsweise für eine Vielzahl von Endbenutzern gleichzeitig diese Filterfunktion realisiert. Das 20 Ausfiltern besteht im wesentlichen aus einer Untersuchung der Daten hinsichtlich ihrer Integrität bezüglich der vorbestimmten, indizierten Themen, dem Qualifizieren der untersuchten Daten hinsichtlich dieser Integrität, dem Speichern dieser Daten und/oder deren Referenzen in einer 25 Datenbank zusammen mit deren Beurteilungsergebnissen und der Entscheidung zwischen Abblocken oder Freigabe.

Wenn eine Firewall die Schnittstelle zu dem Informationsnetzwerk darstellt, kann diese in vorteilhafter Weise

WO 01/90932

PCT/EP01/05746

- 5 -

auch derart erweitert sein, daß die Entscheidung auf Ab-
blocken oder Freigabe von ihr erstellt wird. Dies hat den
Vorteil, daß Wartung und Pflege der Kriterien an einer
einzigen Stelle, einfach für eine Mehrzahl von Endbenut-
zern durchgeführt werden kann, wobei dies gleich in Kom-
bination mit der im Stand der Technik vorhandenen Praxis
der automatischen Virenkontrolle kombiniert werden kann.
5 Im Falle der Nutzung des erfindungsgemäßen Verfahrens
kann das erfindungsgemäße Filterverfahren auch entweder
10 zwangsgesteuert oder vom Endbenutzer freiwillig steuerbar
von seinem entfernt liegenden PC aus angestoßen werden.
In vorteilhafter Weise kann die vorerwähnte Filterfunkti-
on auch durch Computersysteme realisiert werden, die bei
dem vom Endbenutzer gewählten Netzprovider eingesetzt
15 werden.

Die der vorliegenden Erfindung zugrundeliegende Idee be-
steht darin, daß alle angeforderten Netzwerkinformations-
ressourcen, z.B. Webseiten aus dem worldwide web, vor dem
20 Konsum beim Endbenutzer auf ihre Integrität hin kontrol-
liert werden. Dies kann zeitnah zur Anforderung gesche-
hen, sollte aber zeitlich vorzugsweise von der Benutzer-
anforderung entkoppelt werden. Die anforderbaren Informa-
tionsressourcen werden im wesentlichen vollautomatisch
25 durch ein erfindungsgemäßes Filterprogramm untersucht und
qualifiziert. Diese Untersuchung läuft vorzugsweise steu-
erbar in verschiedenen Ebenen ab: Einerseits ist es sinn-
voll, eine 'schwarze Liste' mit Referenzen auf Adressen
bestimmter Datenquellen zu führen, die bekanntermaßen un-

WO 01/90932

PCT/EP01/05746

- 6 -

gewünschte Inhalte zur Verfügung stellen. Andererseits ist es sinnvoll, eine 'grüne Liste' mit Referenzen auf Adressen bestimmter Datenquellen zu führen, die bekanntermaßen gewünschte Inhalte zur Verfügung stellen.

5

Dies können beispielsweise die IP-Adressen der zugehörigen Webserver oder spezielle Webseiten-URLs sein, wenn als Informationsressource das Internet gilt. Diese einfache Form des Filterns benötigt dann lediglich einen Abgleich zwischen der vom Endbenutzer angeforderten Referenz mit den entsprechenden Listen.

Das erfindungsgemäß vorgeschlagene Verfahren ist jedoch wesentlich flexibler und wirksamer als ein solcher passaler Abgleich: denn bestimmte Elemente der aus dem Netzwerk stammenden Daten können einzeln identifiziert und mit einer Wichtung belegt abgespeichert werden. Dabei deckt die Wichtung die verschiedenen, oben genannten, indizierten Themenbereiche, wie zum Beispiel Gewalt, Porno, Sex and Crime, ab. Ziel ist es dabei, möglichst zuverlässig signifikante Informationen aus dem downloadeten Inhalt zu finden, die möglichst eindeutig den Schluß zulassen, daß der betreffende Inhalt als unerwünscht eingestuft werden kann. Ein nackter Hintern kann beispielsweise mit einem Wichtungsprozentsatz von 80 % für den Themenbereich Sex, mit 30 % für den Themenbereich Hardcore, mit 40 % für den Themenbereich Kinderporno, mit 0 % für Themenbereich Gewalt, ebenso 0 % für Werbung belegt werden. Wird beispielsweise ein nacktes Knie als Bildelement

WO 01/90932

PCT/EP01/05746

- 7 -

identifiziert, so bekommt es nur einen Wichtungsprozentsatz von 20 % für Sex und entsprechend reduzierte Prozentsätze für die anderen Themenbereiche.

5 Die von dem Filterprogramm benutzten Suchalgorithmen können dabei auf im Stand der Technik vorhandenen Scan-Funktionen für Bilder, Text und Sprache basieren. In vorteilhafter Weise können die entsprechenden Module direkt von dem erfindungsgemäßen Filterprogramm übernommen und

10 aufgerufen werden. Es werden also für alle Informationskategorien wie beispielsweise Text, Bild, Audio und Video spezifische Elementelisten erzeugt, die jeweils eine möglichst vollständige Auflistung solcher Elemente beinhalten, die einen Beitrag dazu liefern könnten, eine Webseite, oder andere Informationen, die das zugehörige Element enthalten, als unerwünscht oder aber als tolerabel einzustufen. In der Textliste können beispielsweise vulgäre Ausdrücke in verschiedenen Sprachen stehen. In der Bildliste können beispielsweise eine Vielzahl von Ganzkörper-

15 Nacktdarstellungen und die Darstellung einzelner, vorzugsweise nackter oder spärlich bekleideter Körperteile abgespeichert werden. In der Audio-Liste können verschiedene Geräusche gespeichert werden, die für einen oder mehrere der oben genannten Themenbereiche signifikant

20 sind, beispielsweise Stöhnen. Einzelne (Stand-) Bilder bzw. Bildsequenzen von Videos können mit der o.g. Bildliste bearbeitet werden.

25

WO 01/90932

PCT/EP01/05746

- 8 -

Die in dem Filterserver laufende Software sucht vorzugsweise automatisch in einem Grundmodus ständig das Internet ab, um möglichst aktuelle Informationen zu haben.

5 Auch im Stand der Technik vorhandene Techniken der Erkennung von Werbung können hierbei sinnvoll zum Einsatz kommen. Bestimmte, häufig wiederkehrende Motive aller Kategorien können in vorteilhafter Weise als besonders typisch für die entsprechende Kategorie abgespeichert werden, um einerseits als Vergleichsmuster für ähnliche Muster zu dienen und andererseits, um sie bei einem wiederholten Auftreten schneller erkennen zu können, wodurch die Performance beim automatischen Absuchen der Webseiten gesteigert wird.

15 Je nach verfügbarem Speicherplatz und anderen Randbedingungen, beispielsweise den bevorzugten Ziel-Webseiten der angeschlossenen Endbenutzer werden entweder komplette Webseiten auf dem Filterserver selbst gespeichert oder nur Referenzen auf diese Daten, beispielsweise die URL 20 oder andere, signifikante Adressangaben, je nach Netzwerktyp oder Informationsquelle.

25 In bevorzugter Weise werden die Qualifizierungsergebnisse für die einzelnen Web-Inhalte in einer oder mehreren Datenbanken systematisch abgelegt. Durch einen Auswertealgorithmus, der diese Informationen liest und noch weitere Informationen aus dem Umfeld der Webseite hinzunimmt, kann dann bezüglich einer vom Endbenutzer angeforderten Webseite programmgesteuert eine Entscheidung getroffen

WO 01/90932

PCT/EP01/05746

- 9 -

werden, ob die angeforderte Information als unerwünscht oder als tolerabel zu beurteilen ist. Ist sie unerwünscht, so wird die Anforderung nicht bedient und gegebenenfalls ein Hinweis an den Endbenutzer ausgegeben. An- 5 dernfalls wird die Information an den Benutzer ausgegeben.

In den Unteransprüchen finden sich vorteilhafte Weiter- 10 bildungen und Verbesserungen des jeweiligen Gegenstandes der Erfindung.

Gemäß einer bevorzugten Weiterbildung besteht das erfindungsgemäß vorgeschlagene System aus zwei Elementen, zum einen mit einer erweiterten Firewall als Front-End und 15 einem sogenannten Clean Surf Server, im folgenden CSS abgekürzt, als Back-End. Dabei befindet sich das Front-End benutzerseitig auf dem Rechner, über den der Benutzer in das Internet kommt, sei es als Einzelplatz oder in einem lokalen Netzwerk. Der CSS ist in vorteilhafter Weise dem 20 Netzprovider zugeordnet, wo er eine Vielzahl von Endbe- nutzern bedienen kann.

Gemäß einer bevorzugten Weiterbildung läuft das erfindungsgemäße Filterverfahren wenigstens zweistufig ab: In 25 einer Vorstufe auf der Benutzerseite findet ein Abgleich zwischen benutzerangeforderten Inhalten mit als unerwünscht markierten und benutzerseitig gespeicherten Referenzen auf indizierte bzw. nicht-indizierte Inhalte statt. Bei einem Treffer in dieser 'ersten schwarzen Li-

WO 01/90932

PCT/EP01/05746

- 10 -

ste' wird die Benutzeranforderung sofort abgelehnt, ohne den CCS zu kontaktieren. Wenn kein Treffer erfolgt und die URL noch nicht untersucht wurde, wird die Benutzeranforderung an den CCS zur Weiterbearbeitung weiterge-
5 reicht. Ein flexibles, effizientes Verfahren folgt dar-
aus, das vom Einzelplatz-PC bis hin zu großen Endbenut-
zer-Netzwerken skalierbar ist.

Gemäß einer bevorzugten Weiterbildung sind diese Referen-
10 zen Adressen bestimmter Datenquellen, die bekanntermaßen ungewünschte oder gewünschte Inhalte zur Verfügung stel-
len, also beispielsweise IP-Server-Adressen oder URL-
Seiten-Adressierungen. Dies steigert die Performance, da
die Antwortzeit auf die Benutzeranforderung hin sehr ge-
15 ring ist.

Gemäß einer bevorzugten Weiterbildung können die Daten
Webseiten aus einem Netzwerk, insbesondere dem Internet
und dem WorldWideWeb entsprechen. Diese Maßnahme trifft
20 dann den derzeitigen Hauptanwendungsfall der vorliegenden
Erfindung.

Gemäß einer bevorzugten Weiterbildung umfassen die Index-
themen wenigstens eines der folgenden Themen: Sex, Hard-
25 core, Kinderporno, Gewalt, Werbung und Kreditkarten, und
ebenso die Eingabemöglichkeit von Kreditkartennummern.
Eine solche Ausgestaltung unerwünschter Themen bietet ei-
nen relativ objektiven moralischen Maßstab zum 'Sauber-
halten' der dem Benutzer dargebotenen Inhalte.

WO 01/90932

PCT/EP01/05746

- 11 -

Gemäß einer bevorzugten Weiterbildung werden Kombinationen bestimmter Elemente mit für die Kombination typischen Wichtigswerten belegt und abgespeichert. Werden beispielsweise in einem heruntergeladenen Bild sowohl eine Peitsche als auch menschliche Nacktdarstellungen gefunden, so kann mit relativ großer Wahrscheinlichkeit darauf geschlossen werden, daß es sich um eine Webseite mit Sado-/Maso-Inhalten handelt. Daher bekommt die Kombination

5 Peitsche plus nackter Körperteil oder nackter Körper einen hohen Wichtigsfaktor im Themebereich Sex and Crime beziehungsweise Hardcore. Eine Peitsche allein als erkanntes Bildelement würde noch nicht ausreichen, um die Webseite derart einzustufen, da Peitschen ja auch als Reitzubehör dienen. Auch nackte Körperteile beziehungsweise nackte oder spärlich bekleidete Ganzkörperdarstellungen rechtfertigen für sich gesehen noch nicht eine solche Klassifizierung, denn sie sind durchaus üblich bei der Präsentation von Bademode oder Damen- oder Herren-

10

15

20 Unterbekleidung.

In vorteilhafter Weise können auch Kombinationen gebildet werden aus Elementen, die aus verschiedenen Kategorien stammen. Beispielsweise könnte die Kombination 'Stöhnen plus Nacktdarstellung' einen hohen Wichtigsfaktor für den Themebereich Sex. Kämen dann noch Handschellen oder bestimmte andere, bekannte Sado-/Maso-Zubehörteile als weiteres Kombinationselement hinzu, so würde der Wichtigsfaktor für Sex wieder reduziert, der von Hardcore

25

WO 01/90932

PCT/EP01/05746

- 12 -

beziehungsweise Sex and Crime jedoch von 0 % auf einen sehr hohen Wert erhöht werden.

5 In bevorzugter Weise können durch komplexe Abfragealgorithmen und umfassende Reservoirs an Elementen für jede einzelne Informationskategorie themenbereichsspezifisch verlässliche Filterungsgergebnisse erzielt werden. Dabei versteht sich von selbst, daß in diese Abfragealgorithmen viel empirisches Wissen integriert werden sollte, damit 10 diese eine optimale Wirksamkeit erzielen.

15 Gemäß einer bevorzugten Weiterbildung kann ein von einem Endbenutzer spezifiziertes Benutzerprofil hinsichtlich der Definition ungewünschter Daten zusätzlich bei der Auswertung berücksichtigt werden. Dieses Benutzerprofil kann dann in geeigneter Weise mit den vorgegebenen, 'objektiven' Filterkriterien überlagert werden, um effektive, benutzerbezogene, subjektive Filterkriterien aufzustellen. Diese effektiven Kriterien ersetzen dann die 20 oben beschriebenen, objektiven Kriterien. Diese Maßnahme eignet sich insbesondere, um das erfindungsgemäße Verfahren an verschiedene Altersgruppen der Endbenutzer anzupassen. So kann es für Erwachsene durchaus wünschenswert sein, Sexdarstellungen betrachten zu können. Andererseits 25 sollte dieses Minderjährige verwehrt sein, ebenso wie der Zugang zu Gewaltdarstellungen.

Gemäß einer bevorzugten Weiterbildung enthält das Benutzerprofil eine Wichtung der verschiedenen Indexthemen.

WO 01/90932

PCT/EP01/05746

- 13 -

Dies ermöglicht eine einfach zu realisierende Überlagerung der subjektiven mit den objektiven Filterkriterien.

Gemäß einer bevorzugten Weiterbildung werden vom Benutzer 5 abgehende Suchbegriffe oder Seitenanforderungen untersucht, ob sie den oben angegebenen objektiven Filterkriterien genügen. Insbesondere können solche Suchbegriffe auch gefiltert werden. In einem solchen Fall werden beispielsweise die URL-Angaben, die vom Endbenutzer eingegeben 10 wurden, auf ihre Integrität hin untersucht, bevor das dahinterliegende Bild- oder Tonmaterial untersucht wird. Findet sich die angeforderte URL auf der schwarzen Liste, so wird der Zugriff verweigert und der Benutzer entsprechend informiert. Dies steigert die Performance.

15 Das erfindungsgemäße Verfahren kann in vorteilhafter Weise auch zum eigentlichen Ausfiltern unerwünschter Inhalte in Bildsequenzen oder Tonsequenzen oder in Videos für sich betrachtet herangezogen werden. Die Auswertemethode 20 hängt dabei sowohl von der Darstellungsweise als auch von der Übertragungsart ab. Werden beispielsweise in einer Videoübertragung bei den Bildinformationen nur Änderungen gegenüber dem Bild davor übertragen, so muß das 'Grundbild' auf seine Integrität hin untersucht werden. Wenn 25 allerdings die Änderung ein Bildelement einführt, das auf der Indexliste steht, so kann die Wichtung nach dem oben angegebenen Verfahren vollzogen werden.

WO 01/90932

PCT/EP01/05746

- 14 -

Bei streaming-basierten Übertragungen ist es vorteilhaft, die Übertragung durch einen temporären Pufferspeicher laufen zu lassen, und unerwünschte Teilsequenzen entweder zu entfernen oder durch andere, unproblematische Teilsequenzen zu ersetzen. Dies hat den Vorteil, daß keine langweiligen Lücken bei der Wiedergabe der gestreamten Daten entstehen.

ZEICHNUNGEN

10

Ausführungsbeispiele der Erfindung sind in den Zeichnungen dargestellt und in der nachfolgenden Beschreibung näher erläutert.

15 Es zeigt:

Fig. 1 eine schematische Blockdarstellung mit den wesentlichen technischen Funktionselementen und den wichtigsten Schritten während des Einsatzes des erfinderischen Verfahrens gemäß einem bevorzugten Ausführungsbeispiel.

BESCHREIBUNG DER AUSFÜHRUNGSBEISPIELE

25 Fig. 1 zeigt eine schematische Blockdarstellung mit den wesentlichen technischen Funktionselementen und den wichtigsten Schritten während des Einsatzes des erfinderischen Verfahrens gemäß einem bevorzugten Ausführungsbeispiel.

WO 01/90932

PCT/EP01/05746

- 15 -

Ein unternehmensinternes Netzwerk 10 enthält eine Mehrzahl N von Endbenutzer-PCs, von denen zumindest einige zum Surfen im Internet eingerichtet sind. Für den vorliegenden Fall interessiert nur der oben eingezeichnete User-PC 1. Er ist mit Bezugszeichen 12 versehen. Eine aus dem Stand der Technik bekannte Firewall-Netzwerkkomponente ist nun um einige erfindungsgemäße Funktionen erweitert. Diese Komponente ist mit Bezugszeichen 14 versehen. Die Primärfunktion der Firewall 14 bleibt die Ankopplung des Unternehmensnetzes an ein oder mehrere öffentliche Netze, wobei hier das Internet als Beispieldnetz dargestellt ist, siehe oberer Bereich der Figur.

15 Die Firewall-Komponente 14 ermöglicht eine Datenverbindung auf einer Datenleitung 16 zu einem sogenannten Clean Surf Server 18, der als zwischengeschaltete Station dient, um einen direkten Kontakt zwischen Firewall 14 und 20 Internet zu vermeiden. Dieser Server 18 wird im folgenden auch als CSS abgekürzt und arbeitet im wesentlichen als Filterserver.

Der Filterserver 18 ist mit einem Robot-Mechanismus 20 verbunden, der grundsätzlich unabhängig von einer Benutzeranfrage einen automatischen Zugang zum Internet besitzt und eine Vielzahl der dort angebotenen Inhalte auf Text, Ton, Dateninhalt, beispielsweise einem Vorhandensein von Viren, sowie Audio- oder Videosequenzen von Web-

WO 01/90932

PCT/EP01/05746

- 16 -

seiten untersucht. Dies geschieht über eine separate Datenleitung 22. Dieser Robot-Mechanismus enthält ein im Stand der Technik bekanntes Such-Programm, das nach einem vorgegebenen Netzwerksuchschema Webseiten einschließlich aller darauf befindlicher Links aufsuchen und deren Inhalte herunterladen kann. In vorteilhafter Weise arbeitet der Robot-Mechanismus in einem separat stehenden, leistungsfähigen Computer, der vorzugsweise von der Performance her frei skalierbar ist, um sich an den wachsenden Datenbestand im Internet gut anpassen zu können. Damit der Durchsatz gut ist, sollte die Datenleitung 22 vorzugsweise eine sehr hohe Kapazität besitzen.

Der Robot-Computer 20 ist logisch und physikalisch mit einer Reihe von Datenbanken 24 verbunden, in denen für jede Informationskategorie eine große Anzahl von Suchkriterien gespeichert sind. Es gibt also für die Informationskategorie 'Text' eine Datenbank 24a, für die Kategorie 'Bild' eine Datenbank 24b, eine Audio-Datenbank 24c, eine Video-Datenbank 24d sowie optional eine Viren-Datenbank 24e. In all diesen Datenbanken sind separat für jede Kategorie bestimmte Elemente gespeichert, die jeweils für ein oder mehrere, bestimmte, indizierte Themenbereiche relevant sind, wie es oben beschrieben wurde. Die Zusammenfassung mehrerer getrennter Datenbanken in eine einzige oder in eine niedrigere Anzahl von Datenbanken kann je nach Datenbanktyp und gewünschter Performance durchgeführt werden.

WO 01/90932

PCT/EP01/05746

- 17 -

Der Robot-Mechanismus ist weiterhin logisch mit zwei Datenbanken 26 und 28 verbunden. Die Einheiten 20, 24, 26 und 28 bilden zusammen eine funktionsfähige Untereinheit 30, die im Normalfall asynchron vom Filterserver 18 arbeitet und laufend das Internet nach neuen Inhalten hin untersucht, wobei in nicht separat dargestellten Suchverarbeitungs-Servern die gefundenen Webseiten mit den in den Datenbanken 24a, .. 24e gespeicherten Suchkriterien nach unerwünschten Inhalten durchsucht werden. Die Suchergebnisse werden dann in den beiden Datenbanken 26 und 28 abgelegt. Vorzugsweise werden gefundene Einzelelemente zusammen mit einer für sie typischen Wichtung in einer der beiden Datenbanken 26 oder 28 abgelegt.

15 Die Datenbank 26 enthält vorzugsweise die IP-Adressen bestimmter Webserver, die verbotene / unerwünschte Inhalte anbieten. Die Datenbank 28 enthält vorzugsweise verbotene /unerwünschte HTML-Seiten beziehungsweise solche HTML-Seiten, die wenigstens zum Teil unerwünschte Inhalte enthalten, sowie eine entsprechende Klassifizierung.

20 Wird beim automatischen Absuchen durch den Robot-Mechanismus beispielsweise eine HTML-Seite gefunden, die noch nicht in der Datenbank 28 abgespeichert ist, und die noch nicht auf ihre Integrität hin untersucht worden ist, so wird sie dem erfindungsgemäßen Untersuchungsverfahren unterworfen. Die gefundene HTML-Seite möge nun Textinformationen, Bild- und Audio-Informationen enthalten.

WO 01/90932

PCT/EP01/05746

- 18 -

Vorzugsweise parallel zueinander werden nun verschiedene Prozesse gestartet: ein Text-Scan-Prozeß, ein Bildelemente-Scan-Prozeß und ein Audio-Elemente-Scan-Prozeß. Jeder der drei Prozesse isoliert nun, sofern möglich, einzelne Elemente in seiner jeweiligen Kategorie und vergleicht sie mit den in den Datenbanken 24 gespeicherten Suchkriterien. Als Textelement wird nun der Text-String 'Ficken' gefunden. Gleichzeitig findet der Bild-Suchprozeß eine einzeln identifizierte Darstellung, die einer in der Bilddatenbank 24b gespeicherten pornographischen Darstellung sehr ähnlich ist und eine pornographische Pose enthält. Des weiteren trifft der Audio-Suchprozeß auf ein Klangmuster, das sehr große Ähnlichkeit mit einem Klangmuster aufweist, der in der Audio-Datenbank 24c als typisches 'Stöhnen' abgespeichert ist. Jedes gefundene Element wird nun zusammen mit je einem Wichtungsfaktor für jeden der indizierten Themenbereiche in der entsprechenden Datenbank, hier der HTML-Datenbank 28 abgespeichert. Diese Verfahrensweise ermöglicht eine nachträgliche Änderung der Beurteilung durch Korrektur der Wichtungsfaktoren, wenn sich herausstellt, daß ein solcher Korrekturbedarf besteht. Eine solche nachträgliche Änderung kann dann erfolgen, ohne daß alle Seiten und Elemente neu untersucht werden müssen.

Auf der untersuchten Webseite werden als signifikante Elemente also der Text-String 'Ficken', das Audio-Klangmuster eines Stöhnens sowie eine einzige pornographische Pose gefunden.

WO 01/90932

PCT/EP01/05746

- 19 -

Der Text-String 'Ficken' bekommt beispielsweise folgende Wichtungsfaktoren zugeteilt: Sex: 100 %, Hardcore: 50 %, Kinderporno: 40 %, Gewalt: 10 %, Werbung: 0 %, Kreditkarte: 0 %, da keine Eingabemöglichkeit für eine Kreditkartennummer gefunden wurde.

Das Audio-Muster 'Stöhnen' bekommt in der Kategorie Sex 100 %, bei Hardcore 60 %, bei Kinderporno 30 %, bei Gewalt 10 %, bei Werbung 0 % und bei Kreditkarte ebenfalls 0 %.

Die pornographische Pose wird als Bildelement ebenfalls abgespeichert, wobei beispielsweise folgende Wichtungsfaktoren vergeben werden: Sex: 80 %, Hardcore: 30 %, Kinderporno: 40 %, Gewalt: 0 %, Werbung: 0 % und Kreditkarte ebenfalls 0 %.

Nach Bewertung der einzelnen Elemente liest ein komplexer Auswertealgorithmus die gespeicherten Wichtungsprofile und faßt sie zu einer Synthese zusammen, wobei vorzugsweise auch besondere Kombinationen einzelner Textelemente, wie es weiter oben erwähnt wurde, in besonderem Maße berücksichtigt werden.

Werden auf einer Webseite beispielsweise aber nur solche Elemente gefunden, die für sich gesehen und auch in Kombination miteinander keine eindeutigen Schlüsse zulassen, so kann auch die Umgebung der Webseite in die Wichtung

WO 01/90932

PCT/EP01/05746

- 20 -

eingehen: wenn in der hierarchischen Gliederung der Webseite weiter oben schon pornographische Inhalte gefunden wurden, oder wenn die URL der Webseite als pornographisch bekannt gilt, dann wird die Seite ebenfalls als unerwünscht abgeblockt, denn auf Pornoseiten finden sich mit einer hohen Wahrscheinlichkeit ausschließlich pornographische Abbildungen. Ein weiteres Indiz für eine Pornoseite sind Links auf bereits als Pornoseite erkannte Webseiten. Denn auch hier gibt es dann eine hohe Wahrscheinlichkeit, pornographisches Material zu finden.

Auch die IP-Adresse des Webservers könnte herangezogen werden, um von vornherein Webseiten auszuschließen oder um im Zweifelsfall Webseiten auszuschließen. Denn häufig liegen auf Webservern Webseiten, die jeweils einem einzigen Themenbereich aus den verbotenen Themen gewidmet sind. Dieser Fall tritt häufig bei illegalen Darstellungen, wie etwa Kinderporno oder rechtsradikalen, gewaltverherrlichenden Inhalten auf.

20 Darüber hinaus können im Zweifelsfall auch Menschen zur Beurteilung einer Webseite herangezogen werden.

Der Auswertealgorismus kumuliert vorzugsweise die Wichtungsfaktoren aller auf einer Webseite gefundenen Elemente kategorieweise geordnet durch Multiplikation. Wenn beispielsweise fünf Elemente der Kategorie Gewalt gefunden werden, die die Wichtungsfaktoren 90 %, 80 %, 95 %, 75 % und 40 % aufweisen, so werden die Prozentzahlen mul-

WO 01/90932

PCT/EP01/05746

- 21 -

tipliziert, um ein Zwischenergebnis zu bilden. Dies wäre im vorliegenden Fall ein kumulativer Prozentsatz von 0,2052. Dieser wäre bereits als relativ hoch anzusehen, so daß die betroffene Seite als unerwünscht für das weitere Bearbeiten markiert wird.

Finden sich beispielsweise auf einer Webseite fünf Elemente mit den einzelnen Wichtungsfaktoren von 20 %, 15 %, 10 40 %, 50 % und 30 %, also einer weit weniger verbots- trächtigen Elementesammlung, so ergibt sich ein kumulativer Prozentsatz von 0,0018. Der kumulative Prozentsatz liegt also bei der gleichen Anzahl von verwerteten Elementen um etwa zwei Zehnerpotenzen niedriger. Er würde daher nicht als unerwünscht markiert werden, sofern nicht 15 andere Ausnahmetatbestände doch dafür sprechen. Es ist offensichtlich, daß der Auswertealgorithmus die Anzahl der kumulierten Elemente bei seiner Beurteilung gebührend berücksichtigt, denn jeder Wichtungsfaktor, der kleiner als 1 ist, drückt den kumulativen Prozentsatz herunter. 20 Daher kann beispielsweise durch Multiplizieren mit der Anzahl der kumulierten Elemente auf einfache Weise dafür eine Kompensation gefunden werden. Damit ergäbe sich beispielsweise für eine Webseite, die 5 Elemente mit einem jeweiligen Wichtungsfaktor von 90 % aufweist, ein kumulativer Prozentsatz von 0,59, der dann mit 5 multipliziert 25 einen Wert von etwa 3 ergäbe. Bei 10 gefundenen Elementen mit einem solchen Wichtungsfaktor ergäbe sich ein Wert von etwa 3,5, was die Ergebnisse gut vergleichbar macht.

WO 01/90932

PCT/EP01/05746

- 22 -

Allgemein kann auch ein Bewertungsfaktor B durch die Formel

$$B = \text{Summe}(p_i E_i)/n$$

5

gewonnen werden, wobei p_i die Wichtungsfaktoren darstellen, E_i die Elemente und n die Anzahl der Elemente.

Nach einer gewissen Vorlaufzeit, während der das Robot-
10 system das Internet durchsucht hat bzw. auf bereits vor-
handene Datenbanken zurückgreift und die Beurteilungser-
gebnisse in den beiden Datenbanken 26 und 28 abgespei-
chert hat, kann ein gewisser Teil der von dem User-PC 12
gemachten Anfragen an HTML-Seiten unter Berücksichtigung
15 der Beurteilungsergebnisse bearbeitet werden. Dazu wird
wie folgt vorgegangen: Die bereits untersuchten Webseiten
werden als untersucht markiert. Der Endbenutzer am User-
PC 12 definiert eine Anfrage nach einer bestimmten HTML-
Seite auf seinem PC, indem er in einem Browser eine be-
20 stimmte Aktion durchführt, wie es durch Anklicken eines
Links oder Eingabe einer URL der Fall sein kann.

Wenn die angeforderte Adresse noch nicht lokal als er-
wünscht oder unerwünscht klassifiziert wurde, wird die
25 Anfrage wird nun in einem separaten Prozeß der erweiter-
ten Firewall 14 bearbeitet und über die Leitung 16 zu-
nächst an den Clean Surf Server CSS 18 weitergeleitet,
der seinerseits die weitere Kontrolle bei der Bearbeitung
übernimmt.

WO 01/90932

PCT/EP01/05746

- 23 -

1. Zunächst wird festgestellt, ob die aktuell angeforderte Webseite bereits untersucht wurde oder nicht. Falls nicht, wird sie aus dem Internet downgeladen und wie 5 oben beschrieben beurteilt, wobei das Beurteilungsergebnis in der Datenbank 28 für HTML-Seiten abgespeichert wird.

2. Danach, ebenso wie in dem Fall, in dem die angeforderte 10 Webseite bereits vor Absenden der Benutzeranforderung untersucht war, wird festgestellt, ob sie als unerwünscht gilt oder nicht. Dies kann durch Setzen eines Flags in dem entsprechenden HTML-Datensatz und Abfragen dieses Flags erfolgen. Je nach Untersuchungsergebnis kann dann 15 der Zugang zur angeforderten Webseite ermöglicht oder abgeblockt werden.

3. Obwohl die vorliegende Erfindung anhand eines bevorzugten Ausführungsbeispiels vorstehend beschrieben wurde, ist 20 sie darauf nicht beschränkt, sondern auf vielfältige Weise modifizierbar.

4. So kann beispielsweise der Clean Surf Server im Falle eines Abblockens an die Firewall zurückmelden, warum diese 25 Seite nicht freigegeben worden ist, wobei die Firewall ab einer einstellbaren Häufigkeit von Anforderungen den Systemadministrator automatisch benachrichtigen kann, daß eine bestimmte Webseite in einem bestimmten Zeitintervall relativ häufig angefordert wurde. Weiterhin kann fest-

WO 01/90932

PCT/EP01/05746

- 24 -

gehalten werden, welche Webseiten angefordert wurde, welche Suchbegriffe verwendet werden, wieviele Verweigerungen es gab, etc.

- 5 Der Systemadministrator kann dann Maßnahmen ergreifen, um im Bedarfsfall die Webseite doch freizugeben oder, falls dies nicht beabsichtigt ist, andere Maßnahmen treffen, je nach Art der Webseite.
- 10 Das erfindungsgemäße Programm kann in vielerlei Ausgestaltungen installiert werden. Vorteilhaft ist eine spezielle Software oder Netzwerkkarte, auf die nur über ein geschütztes Paßwort zugegriffen werden kann, damit der Endbenutzer z.B. den Standard-Gateway oder Proxiserver
- 15 von sich aus nicht umstellen kann. Dies kann auch als Kindersicherung dienen.

In vorteilhafter Weise können einzelne der vorhandenen Programmfunctionen des erfinderischen Verfahrens auch in 20 einen herkömmlichen Web-Browser integriert sein.

Des weiteren kann in einer unter Umständen abgespeckten Version des erfinderischen Verfahrens die Funktion von Clean Surf Server 18 und Firewall 14 vollständig auf den 25 End-User-PC gebracht werden, indem beispielsweise ein Verzeichnis aller nicht-erwünschten Inhalte, gekennzeichnet etwa durch die URLs oder die IP-Adresse von Webser-vern abgefragt wird, bevor eine Benutzeranforderung dem Endbenutzer zugänglich gemacht wird. Eine solche 'schwar-

WO 01/90932

PCT/EP01/05746

- 25 -

ze Liste' kann beispielsweise auch in Form einer CD einzeln verkauft oder über das Internet oder sonstige mögliche Datenübertragungen downloadbar sein.

5 Des weiteren besteht die Möglichkeit, daß ein Endbenutzer, wenn er trotz Filterung eine unerwünschte Seite erhält, dies dem CSS rückmeldet, beispielsweise durch Betätigen eines eigenen Buttons in dem von ihm benutzten Browser.

10 Auch ein Bonussystem kann für verschiedene Zwecke in Kombination mit bestimmten der vorerwähnten Merkmale implementiert werden. Des weiteren können in einer speziellen Weiterbildung des erfinderischen Verfahrens solche Webs Seiten oder allgemeine Inhalte, die als hochgradig unerwünscht beurteilt wurden, automatisch einer separaten Behandlung unterzogen werden, die beispielsweise das Informieren einer zuständigen Behörde miteinschließt.

20 Mit der vorgeschlagenen Firewall-Erweiterung können ganze Netze ebenso wie Einzelrechner abgesichert werden. Das erfinderische Konzept ist nicht beschränkt auf das Absuchen verbotener Inhalte im Internet oder WorldWideWeb. Auch andere Netzwerke, wie beispielsweise Intranets können durchsucht werden.

Des weiteren können sogenannte Pushings und Pop-Ups, also ein automatisches Aufdrücken von Seiten beziehungsweise ein automatisches Aufmachen von Fenstern verhindert werden.

WO 01/90932

PCT/EP01/05746

- 26 -

den. Des weiteren können vorhandene Technologien wie Tunneling, also eine virtuelle Netzwerkprotokollverschachtelung mit implementiert werden. Auch kann die erweiterte Firewall und die CSS-Komponente auf einem Rechner oder 5 System implementiert sein, der entfernt vom Endbenutzer-PC liegt und optional ebenfalls noch als Web-Server dient.

Des weiteren ist es möglich, durch Triggering-Mechanismen 10 vom Stand der Technik zahlreiche andere Aktionen auszulösen, wenn eine Webseite als ungewünscht markiert wurde. So kann es beispielsweise sinnvoll sein, den Web-Master der zuständigen Seite zu informieren, beispielsweise durch automatisches Versenden einer eMail. Der Web-Master 15 hat dann die Möglichkeit, Stellung zu nehmen oder die Seite möglicherweise zu verändern.

WO 01/90932

PCT/EP01/05746

27

Verfahren und Vorrichtungen zum Abblocken von aus einem
Netzwerk anforderbaren Daten

5

BEZUGSZEICHENLISTE:

10	Netzwerk (LAN)
12	Endbenutzer-PC
14	Firewall
16	Datenleitung
18	Clean Surf Server (CSS)
20	Robot-Mechanismus
22	separate Datenleitung
24-28	Datenbanken
30	Untereinheit

10

WO 01/90932

PCT/EP01/05746

28

5

PATENTANSPRÜCHE

- 10 1. Verfahren zum Abblocken von aus einem Netzwerk anforderbaren Daten mit unerwünschtem Inhalt, enthaltend die Schritte,
Daten über einen vorbestimmte Filterkriterien verwenden den Clean Surf Server (CCS) (18) als Filterserver zwischen einem Endbenutzer-Computer (12) und dem Netzwerk aus diesem anzufordern,
um unerwünschte Daten von zu tolerierenden Daten zu unterscheiden.
- 20 2. Verfahren nach Anspruch 1, verwendet in einem Firewallsystem (14), um den Empfang unerwünschter Inhalte an mehreren, miteinander vernetzten Computern (10) zu verhindern.
- 25 3. Verfahren nach einem der vorstehenden Ansprüche, den weiteren Schritt enthaltend, in einer benutzerseitigen Vorstufe einen Abgleich zwischen benutzerangeforderten Inhalten mit als unerwünscht oder erwünscht markierten und benutzerseitig gespeicherten Referenzen durchzufüh-

WO 01/90932

PCT/EP01/05746

29

ren,

bei einem Treffer die Benutzeranforderung abzulehnen, und
5 andernfalls die Anforderung an den CCS zur Weiterbearbei-
tung weiterzugeben.

4. Verfahren nach dem vorstehenden Anspruch, wobei die
Referenzen Adressen bestimmter Datenquellen, die bekann-
10 termaßen ungewünschte oder gewünschte überprüfbare Inhal-
te zur Verfügung stellen, als Serveradressen oder als
Seitenadressen enthalten.

5. Verfahren zum Ausfiltern von aus einem Netzwerk anfor-
15 derbaren Daten mit unerwünschtem Inhalt, enthaltend die
Schritte,

Untersuchen der Daten hinsichtlich ihrer Erwünschtheit,
20 Qualifizieren der untersuchten Daten hinsichtlich ihrer
Erwünschtheit,

Speichern von Netzwerkdaten und/ oder deren Referenzen in
einer Datenbank zusammen mit deren Beurteilungsergebnissen
25 hinsichtlich unerwünschtem Inhalt,

Vorenthalten oder Freigeben von Benutzeranforderungen auf
diese Daten je nach Maß ihrer Erwünschtheit.

WO 01/90932

PCT/EP01/05746

30

6. Verfahren nach Anspruch 5, wobei die Daten Webseiten aus einem Netzwerk, insbesondere dem Internet entsprechen.

5 7. Verfahren nach Anspruch 5 oder 6, wobei die Daten wenigstens auf eines von Text, Bild, Ton, oder Virus-Befallenheit untersucht werden, und die Daten bezüglich ihrer Zugehörigkeit zu verschiedenen Indexthemen mit einer Wichtung beurteilt werden.

10 8. Verfahren nach einem der vorstehenden Ansprüche 5 bis 7, wobei die Indexthemen wenigstens eines von Sex, Hardcore, Kinderporno, Gewalt, Werbung, Eingabemöglichkeit von Kreditkartennummern umfassen.

15 9. Verfahren nach einem der Ansprüche 1 bis 8, wobei bestimmte Elemente der Daten einzeln identifiziert und mit einer Wichtung belegt abgespeichert werden.

20 10. Verfahren nach einem der Ansprüche 1 bis 9, wobei Kombinationen bestimmter Elemente mit für sie typischen Wichtungswerten belegt abgespeichert werden.

25 11. Verfahren nach einem der vorstehenden Ansprüche, weiter enthaltend den Schritt,

Auswerten eines Benutzerprofils hinsichtlich der Definition ungewünschter Daten,

WO 01/90932

PCT/EP01/05746

31

Überlagern des Benutzerprofils mit dem Erwünschtheitsprofil zur Bestimmung eines subjektiven Erwünschtheitsprofils, um individuellen Filterkriterien zu genügen,

5

Vorenthalten oder Freigeben von Benutzeranforderungen auf diese Daten je nach Maß ihrer subjektiven Erwünschtheit,

Übermitteln einer Begründung im Falle des Vorenthaltens
10 an den Benutzer.

12. Verfahren nach dem vorstehenden Anspruch, wobei das Benutzerprofil eine Wichtung verschiedener Indexthemen enthält

15

13. Verfahren nach einem der vorstehenden Ansprüche, wobei vom Benutzer abgehende Suchbegriffe oder Seitenanforderungen hinsichtlich Anforderungen unerwünschten Inhalts untersucht werden und optionellerweise weitergemeldet werden.

20

14. Verwendung des Verfahrens nach einem der Ansprüche 5 bis 14 zum Ausfiltern unerwünschter Inhalte von Bildsequenzen oder Tonsequenzen oder Videos.

25

15. Verfahren nach dem vorstehenden Anspruch, wobei die Übertragung bei Streaming-basierten Übertragungen gepuffert verläuft und unerwünschte Teilsequenzen entfernt

WO 01/90932

PCT/EP01/05746

32

oder durch andere Teilsequenzen ersetzt werden.

16. Computerprogramm enthaltend Codeabschnitte zur Ausführung von Schritten des Verfahrens nach einem der Ansprüche 1 bis 4 oder 5 bis 15.

17. Computerprogrammerzeugnis, gespeichert auf einem computerlesbaren Datenträger, enthaltend computerlesbare Programmeinrichtungen, um einen Computer zur Ausführung von Schritten des Verfahrens nach einem der Ansprüche 1 bis 4 oder 5 bis 15 zu veranlassen, wenn es in eine Computer geladen wird.

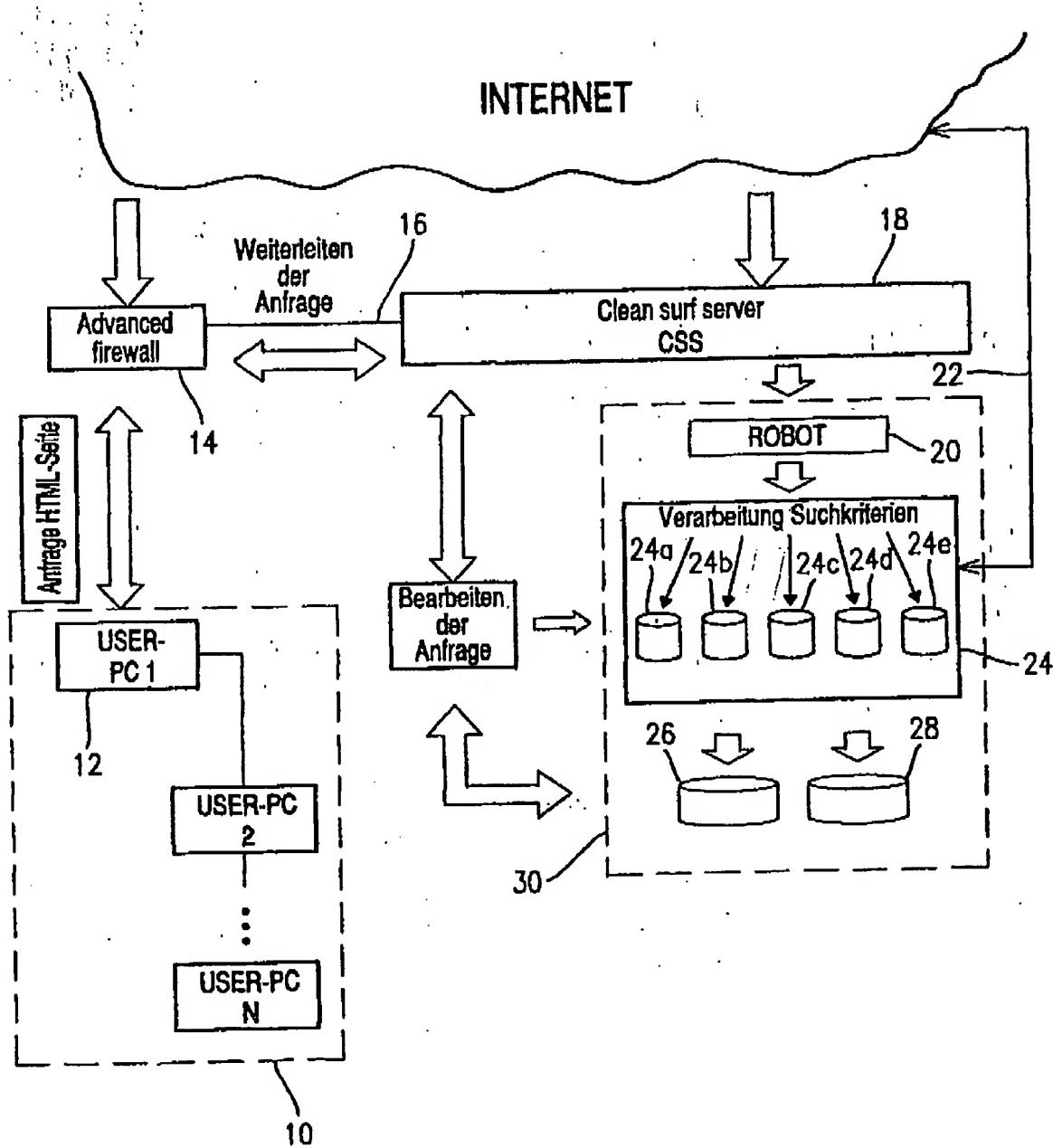
18. Computersystem, enthaltend Mittel zur Durchführung des Verfahrens nach einem der Ansprüche 1 bis 4 oder 5 bis 15.

19. Computersystem, enthaltend Mittel zur Durchführung des Verfahrens nach einem der Ansprüche 1 bis 4 in Kombination mit dem Verfahren nach Ansprüchen 5 bis 15.

WO 01/90932

PCT/EP01/05746

1/1



(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG(19) Weltorganisation für geistiges Eigentum
Internationales Büro

(43) Internationales Veröffentlichungsdatum
29. November 2001 (29.11.2001)

PCT

(10) Internationale Veröffentlichungsnummer
WO 01/090932 A3(51) Internationale Patentklassifikation⁷: G06F 17/30, 17/60 (74) Anwalt: BARTH, Stephan; Reinhard, Skuhra, Weise & Partner GbR, Postfach 44 01 51, 80750 München (DE).

(21) Internationales Aktenzeichen: PCT/EP01/05746

(81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(22) Internationales Anmeldedatum: 18. Mai 2001 (18.05.2001)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

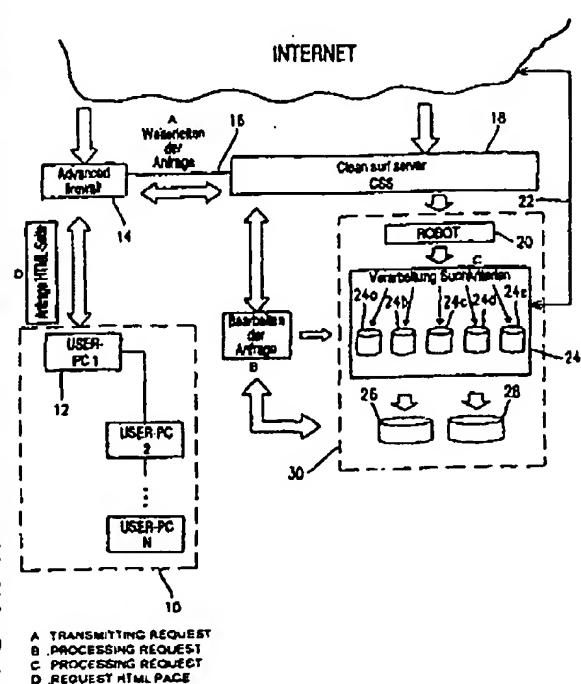
(30) Angaben zur Priorität: 100 24 733.4 19. Mai 2000 (19.05.2000) DE

(84) Bestimmungsstaaten (regional): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Anmelder und
(72) Erfinder: SPEHR, Clemente (DE/DE); Holzstr. 33, 80469 München (DE).*[Fortsetzung auf der nächsten Seite]*

(54) Title: METHOD AND DEVICES FOR BLOCKING DATA WHICH CAN BE REQUESTED FROM A NETWORK

(54) Bezeichnung: VERFAHREN UND VORRICHTUNG ZUM ABBLOCKEN VON AUS EINEM NETZWERK ANFORDERBAREN DATEN

**WO 01/090932 A3**

(57) Abstract: The invention relates to a method and devices for blocking resources which can be requested from a network and which have an undesirable content, and to a method for filtering data of this type out of any quantity of data. Requested web pages are e.g. checked for integrity before being used by the end user. The information resources that can be requested are essentially fully automatically checked and qualified by an inventive filtering program based on the detection and weighted evaluation of individual suspicious information elements.

(57) Zusammenfassung: Ein Verfahren und Vorrichtungen zum Abblocken von aus einem Netzwerk anforderbaren Ressourcen mit unerwünschtem Inhalt, sowie ein Verfahren zum Ausfiltern solcher Daten aus einer beliebig grossen Datenmenge wird vorgeschlagen. Angeforderte Webseiten werden beispielsweise vor dem Verbrauch beim Endbenutzer auf ihre Integrität hin kontrolliert. Die anforderbaren Informationsressourcen werden im wesentlichen vollautomatisch durch ein erfundsgemässes Filterprogramm beruhend auf Erkennung und gewichteter Bewertung einzelner "verdächtiger" Informationselemente untersucht und qualifiziert.

WO 01/090932 A3


Zur Erklärung der Zweibuchstaben-Codes und der anderen
Abkürzungen wird auf die Erklärungen ("Guidance Notes on
Codes and Abbreviations") am Anfang jeder regulären Ausgabe
der PCT-Gazette verwiesen.

Veröffentlicht:

mit internationalem Recherchenbericht

(88) Veröffentlichungsdatum des internationalen

Recherchenberichts: 27. Februar 2003

INTERNATIONAL SEARCH REPORT

In onal Application No
PCT/EP 01/05746A. CLASSIFICATION OF SUBJECT MATTER
IPC 7. G06F17/30 G06F17/60

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 911 043 A (DUFFY ROBERT A ET AL) 8 June 1999 (1999-06-08) the whole document	5-13, 16-19
Y		1-4, 14, 15
X	US 5 996 011 A (HUMES DONALD CREIG) 30 November 1999 (1999-11-30)	1-4, 16-19
A	abstract column 2, line 30 -column 4, line 22	5-13
Y	US 5 696 898 A (GROSSE ERIC ET AL) 9 December 1997 (1997-12-09)	1-4
A	column 1, line 60 -column 3, line 16 column 4, line 27 -column 5, line 5	5-19
		-/-

 Further documents are listed in the continuation of box C. Patent family members are listed in annex.

* Special categories of cited documents :

- *A*: document defining the general state of the art which is not considered to be of particular relevance
- *E*: earlier document but published on or after the International filing date
- *L*: document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O*: document referring to an oral disclosure, use, exhibition or other means
- *P*: document published prior to the International filing date but later than the priority date claimed

T later document published after the International filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

& document member of the same patent family

Date of the actual completion of the International search

Date of mailing of the International search report

15 November 2002

21/11/2002

Name and mailing address of the ISA

European Patent Office, P.O. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+31-70) 340-3016

Authorized officer

Abbing, R

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/EP 01/05746

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5 757 417 A (ARAS CAGLAN M ET AL) 26 May 1998 (1998-05-26) abstract column 2, line 38 -column 3, line 45 column 12, line 34 -column 13, line 42	14,15
A	WO 99 54827 A (ELSWICK PAUL BLAIR ;HUGES PATRICK ALAN (US)) 28 October 1999 (1999-10-28) abstract page 3, line 1 -page 4, line 3	1,2,4-6, 13,16-19

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 01/05746

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 5911043	A	08-06-1999	NONE		
US 5996011	A	30-11-1999	NONE		
US 5696898	A	09-12-1997	CA 2196867 A1 CN 1159234 A EP 0793826 A1 JP 2001526804 T WO 9715008 A1	07-12-1996 10-09-1997 10-09-1997 18-12-2001 24-04-1997	
US 5757417	A	26-05-1998	DE 69619078 D1 DE 69619078 T2 EP 0777385 A2 JP 9238321 A	21-03-2002 17-10-2002 04-06-1997 09-09-1997	
WO 9954827	A	28-10-1999	US 6065055 A AU 3863599 A CA 2328941 A1 EP 1080417 A1 WO 9954827 A1 US 6389472 B1	16-05-2000 08-11-1999 28-10-1999 07-03-2001 28-10-1999 14-05-2002	

Form PCT/ISA/210 (Patent family annex) (July 1992)

INTERNATIONALER RECHERCHENBERICHT

In **nationales Albenzeichen**
PCT/EP 01/05746

A. KLASIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
IPK 7 G06F17/30 G06F17/60

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierte Mindestprästoff (Klassifikationssystem und Klassifikationssymbole)
IPK 7 G06F

Recherchierte aber nicht zum Mindestprästoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der Internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

EPO-Internal

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	US 5 911 043 A (DUFFY ROBERT A ET AL) 8. Juni 1999 (1999-06-08) das ganze Dokument	5-13, 16-19
Y		1-4, 14, 15
X	US 5 996 011 A (HUMES DONALD CREIG) 30. November 1999 (1999-11-30)	1-4, 16-19
A	Zusammenfassung Spalte 2, Zeile 30 -Spalte 4, Zeile 22	5-13
Y	US 5 696 898 A (GROSSE ERIC ET AL) 9. Dezember 1997 (1997-12-09)	1-4 5-19
A	Spalte 1, Zeile 60 -Spalte 3, Zeile 16 Spalte 4, Zeile 27 -Spalte 5, Zeile 5	
		-/-

Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen :

A Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

E Älteres Dokument, das jedoch erst am oder nach dem Internationalen Anmeldedatum veröffentlicht worden ist

L Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungstermin einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgenutzt)

O Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

P Veröffentlichung, die vor dem Internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

T Spätere Veröffentlichung, die nach dem Internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

Y Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erforderlicher Tätigkeit beruhend betrachtet werden

Z Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erforderlicher Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

& Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der Internationalen Recherche

Absendedatum des Internationalen Recherchenberichts

15. November 2002

21/11/2002

Namn und Postanschrift der Internationalen Recherchenbehörde
 Europäisches Patentamt, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel (+31-70) 340-2040, TX 31 861 epo nl.
 Fax (+31-70) 340-3016

Bevollmächtigter Bediensteter

Abbing, R

INTERNATIONALER RECHERCHENBERICHT

Internes Aktenzeichen
PCT/EP 01/05746

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
Y	US 5 757 417 A (ARAS CAGLAN M ET AL) 26. Mai 1998 (1998-05-26) Zusammenfassung Spalte 2, Zeile 38 -Spalte 3, Zeile 45 Spalte 12, Zeile 34 -Spalte 13, Zeile 42	14,15
A	WO 99 54827 A (ELSWICK PAUL BLAIR ;HUGES PATRICK ALAN (US)) 28. Oktober 1999 (1999-10-28) Zusammenfassung Seite 3, Zeile 1 -Seite 4, Zeile 3	1,2,4-6, 13,16-19

Formblatt PCT/GA/210 (Fortsetzung von Blatt 2) (Juli 1992)

INTERNATIONALER RECHERCHENBERICHT

In nationales Aktenzeichen
 PCT/EP 01/05746

Im Recherchenbericht angeführtes Patentdokument		Datum der Veröffentlichung	Mitglied(er) der Patentfamilie		Datum der Veröffentlichung
US 5911043	A	08-06-1999	KEINE		
US 5996011	A	30-11-1999	KEINE		
US 5696898	A	09-12-1997	CA 2196867 A1 CN 1159234 A EP 0793826 A1 JP 2001526804 T WO 9715008 A1		07-12-1996 10-09-1997 10-09-1997 18-12-2001 24-04-1997
US 5757417	A	26-05-1998	DE 69619078 D1 DE 69619078 T2 EP 0777385 A2 JP 9238321 A		21-03-2002 17-10-2002 04-06-1997 09-09-1997
WO 9954827	A	28-10-1999	US 6065055 A AU 3863599 A CA 2328941 A1 EP 1080417 A1 WO 9954827 A1 US 6389472 B1		16-05-2000 08-11-1999 28-10-1999 07-03-2001 28-10-1999 14-05-2002